



GB04/04432



INVESTOR IN PEOPLE

The Patent Office  
Concept House  
Cardiff Road  
Newport  
South Wales  
NP10 8QQ

REC'D. 07 DEC 2004

WIPO

PCT

**PRIORITY  
DOCUMENT**

SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

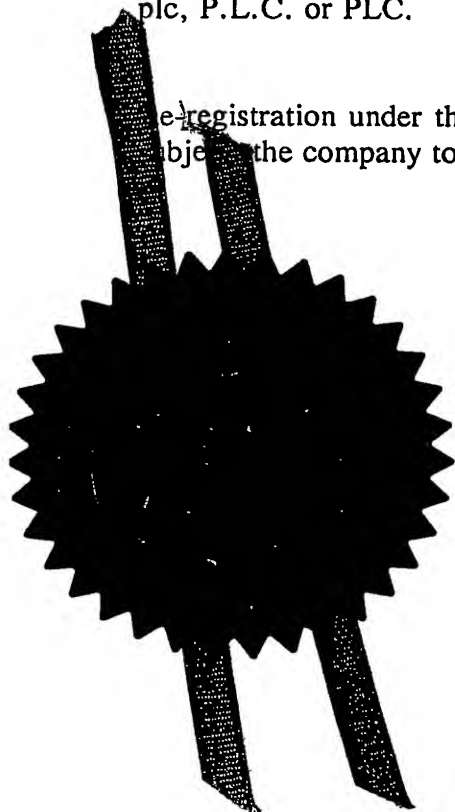
In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed

Dated

24 November 2004



**BEST AVAILABLE COPY**

# Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)



The Patent Office

Cardiff Road  
Newport  
South Wales  
NP9 1RH

1. Your reference P35714GB/KVC

2. Patent application number  
(The Patent Office will fill in this part)

0324456.3

20 OCT 2003

3. Full name, address and postcode of the or of each applicant (underline all surnames)

Isis Innovation Ltd  
Ewert House  
Ewert Place, Summertown  
Oxfordshire OX2 7SG

Patents ADP number (if you know it)

3998564003

If the applicant is a corporate body, give the country/state of its incorporation

UK

4. Title of the invention

Parallel DNA sequencing methods.

5. Name of your agent (if you have one)

"Address for service" in the United Kingdom to which all correspondence should be sent (including the postcode)

Kilburn & Strode  
20 Red Lion Street  
London  
WC1R 4PJ

Patents ADP number (if you know it)

125001

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and (if you know it) the or each application number

Country

Priority application number  
(if you know it)

Date of filing  
(day / month / year)

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application

Date of filing  
(day / month / year)

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? (Answer 'Yes' if:

- a) any applicant named in part 3 is not an inventor, or
  - b) there is an inventor who is not named as an applicant, or
  - c) any named applicant is a corporate body.
- See note (d))

YES

# Patents Form 1/77

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form

Description 37

Claim(s) 6

Abstract --

Drawing(s) --

10. If you are also filing any of the following, state how many against each item.

Priority documents --

Translations of priority documents --

Statement of inventorship and right to grant of a patent (Patents Form 7/77) --

Request for preliminary examination and search (Patents Form 9/77) --

Request for substantive examination (Patents Form 10/77) --

Any other documents (please specify) --

11. I/We request the grant of a patent on the basis of this application.

Signature

Date 20.10.03

Kilbarn and Shado

12. Name and daytime telephone number of person to contact in the United Kingdom  
Kristina Cornish  
Tel: 020 7539 4200

## Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

## Notes

1. If you need help to fill in this form or you have any questions please contact the Patent Office on 0345 20007

2. This form must be filled in by the applicant or their agent.

3. This form must be filled in by the applicant or their agent.

## Parallel DNA sequencing methods

The present invention relates to a method of sequencing a target polynucleotide, a method of comparing polynucleotide sequences and a method of resolving ambiguities in a polynucleotide sequence.

The current methods of sequencing nucleotides are both time consuming and expensive. Sequencing the human genome for the first time took more than ten years and hundreds of millions of dollars. If re-sequencing of genomes and *de novo* sequencing of pathogens and model organisms could be performed several orders of magnitude faster and more cheaply it would open up new avenues for disease genetics and functional genomics.

Sanger dideoxy sequencing which provides sequence information rather indirectly, by looking at the differences in gel-migration of a ladder of terminated extension reactions provided the means to sequence the consensus human genome. Now, however, the need for large scale re-sequencing of individual human genomes and *de novo* sequencing and re-sequencing in pathogens and model organisms require cheaper and faster alternatives to be developed.

Several methods that would avoid gel electrophoresis, cloning or the Polymerase-chain reaction (PCR) have been suggested. One method, "sequencing by synthesis" (SbS), involves the identification of each nucleotide immediately following its incorporation by a polymerase into an extending DNA strand. One SbS approach, pyrosequencing, is widely used for SNP (single-nucleotide polymorphism) typing. In this case, the detection is based on pyrophosphate (PPi) release, its conversion to ATP, and the production of visible light by firefly luciferase. However, because the signal is diffusible, pyrosequencing cannot take advantage of the massive degree of parallelism that becomes available when surface immobilised reactions are analysed.

WO96/27025 discloses a sequencing by synthesis strategy, which involves a removable blocking group. In this method the nucleotide which is incorporated is

modified so that it has a blocking group, which prevents the addition of further nucleotides, and a label. Once the incorporated nucleotide has been identified the blocking group is removed to allow the next nucleotide to be incorporated. The

5 downside of this approach is that the chemistry used to remove the blocking group can damage the DNA and it is difficult to ensure that the deblocking reaction goes to completion which means that sequence acquisition on individual molecules can dephase from each other.

Another strategy removes an entire base rather than just a blocking group.

10 US2003013101 discloses a method of removing a base in its entirety, using the 3' to 5' exonuclease activity of a DNA polymerase. The use of enzymes is gentle on the DNA and reactions can be stoichiometric. However, this method may remove more nucleotides than are added during a specific synthesis cycle, potentially causing errors and repetitions in the sequence.

15

The present invention provides new sequencing by synthesis methods to overcome the shortcomings of the methods disclosed in the prior art.

20 In one aspect the present invention provides a method of sequencing a target polynucleotide comprising the steps of:

- (a) Carrying out template derived nucleotide synthesis utilising a labelled nucleotide;
- (b) detecting the presence or absence of said labelled nucleotide;
- (c) replacing said labelled nucleotide with an unlabelled nucleotide; and
- 25 (d) repeating steps a) to c);

with the proviso that if said labelled nucleotide is labelled with a label directly

“Template derived polynucleotide synthesis” as used herein means forming a polynucleotide molecule utilising a polymerizing reagent that specifically incorporates nucleotides consistent with the well known Watson Crick base pairing rules, using the target nucleotide sequence as a template. The incorporation may be of nucleotide analogues nucleotide mimics or other molecules which can be templated by a polynucleotide and in which pairing is by well defined rules (Eckardt et al 2002; Czapinski, et al, 2001). For example, high-fidelity templating of DNA base shape mimics without forming Watson Crick Bonds has been reported (Delaney et al, 2003). Vice versa the template may be any molecule which can template polynucleotide synthesis.

Polymerizing reagents include DNA polymerases, RNA polymerases, RNA transcriptases, reverse transcriptases, or ligases, as well as chemical reagents that enable template directed polymerization. As used herein “polymerising reagent” also includes molecules or complexes that are capable of enforcing high fidelity base pairing according to well defined rules, regardless of whether they catalyse the addition of a single nucleotide. They can be natural, such as those listed above and ribozymes, or artificial such as abzymes. The polymerizing reagent may comprise one or more chemical reagents. For example, template directed ligation can be mediated by chemical reactions ( Xu et al, 2001; G. von Kiedrowski, 1986).

The target polynucleotide and synthesised polynucleotide can each independently be strands of RNA or DNA. The DNA can be genomic DNA, or cDNA. The RNA can be mRNA, or genomic RNA, such as that from a virus. Alternatively the target polynucleotide and/or synthesised polynucleotide can have an amide backbone formed through peptide nucleic acids (PNA) or a ribose P backbone, as formed by DNA.

The synthesis process can involve annealing a primer to the template polynucleotide. The primer can then be extended by template derived synthesis. The primer consists of 5-100 nucleotides, preferably 10-75, 15-65, 20 – 55, 25 – 50, or 30 - 45 nucleotides. The primer may be labelled. The primer may be made and then

hybridised to the target polynucleotide. Alternatively nicks can be made in double stranded molecules using for example, Deoxyribonuclease 1 (DNase 1) optimised so that the distance between each nick is reasonably defined. The intact strand is the target polynucleotide to which a series of primers are annealed. Synthesis, and thus

5 sequencing can start at each nick site and the non-template strand become progressively displaced. Whether a nick seeds displacement synthesis in a sense or antisense strand is revealed by the direction of migration of the sequencing signal. The template polynucleotide is preferably attached to a solid surface.

10 The template can be attached indirectly to a surface, via a polymerizing reagent which is attached to the surface or it can be captured by a capture probe/primer. The capture may be of a single stranded target or a cohesive termini or "sticky end" of a double stranded template.

15 RNA Promoters native to the template DNA can be used for RNA synthesis by RNA polymerase. Alternatively, extrinsic promoters for specific polymerases can be incorporated by being part of a capture probe or by transposon directed integration into sites along the polynucleotide. T7 and T3 RNA polymerase promoters are preferred extrinsic promoters.

20 The term "labelled nucleotide" as used herein means any of the standard deoxyribonucleotides, or ribonucleotides which is attached to a label. Alternatively the nucleotides include any modified nucleotides or variations with other bases which pair with other bases according to defined rules, such as the Watson-Crick base pair rules. Thus the labelled nucleotide can be a labelled peptide nucleotide capable of forming PNA.

25

sequence information. The random portion stabilises the interaction with the template and provides sites for attachment of the label.

5 The label can be radioactive (such as  $^{32}\text{P}$ ), or more preferably a fluorescent tag. The fluorescent tag may be a dye molecule such as a fluorophore, for example Cy3 or Cy5 (Amersham, UK), ROX (carboxy-x-rhodamine), TAMRA (tetramethylrhodamine), Oregon Green®, Vistra Green™, Fluorescein, PicoGreen®, BODIPY® and Texas Red®. Such fluorophores are commercially available, for example, from Atto-tech (Germany) or Molecular Probes (USA) (Kricka LJ.). Alternatively the label can be a  
10 tag which can be identified due to its physiochemical properties, eg electrophoretic mobility or an electric charge. Alternatively a raman signal can be detected, for example Surface enhanced raman scattering (SERRs) (Kneipp 1999; Zander 2002).

15 The label can be attached directly through a covalent bond to the nucleotide, or via a linkage. The linkage preferably comprises a cleavable bond, for example a photocleavable bond, or a bond which is cleavable by a mild reducing agent, such as a disulfide bridge.

20 The processing of sequencing relies on the base pairing that occurs between nucleotides to form a double stranded polynucleotide molecule, according to the Watson-Crick base pairing rules. At each position in a nucleotide molecule, one of the four nucleotides can be incorporated. The nucleotide incorporated into the extending primer or into an RNA copy is normally the correct base that pairs with the base in the target polynucleotide.

25 The sequencing method can be carried out in two ways. The four nucleotides can be labelled with the same label e.g. one fluorophore. The primer/template polynucleotides can be contacted with one nucleotide (e.g. adenine). The unincorporated nucleotides can then be removed either by being washed away or  
30 degraded by an enzyme such as apyrase. Any nucleotides that have been incorporated can then be detected. This process can then be repeated with the other three nucleotides (e.g. thymine, cytosine and guanine). Alternatively the four nucleotides

can be differentially labelled i.e. each has a different label or fluorophore. In this case the primer and template polynucleotides are contacted with two or more of the labelled nucleotides at the same time. After removal of any free nucleotides, the incorporated bases are detected. The use of four differentially labelled nucleotides  
5 allows real time monitoring of the synthesis process.

In one embodiment the label can be "directly attached" to the nucleotide via a covalent bond to the base, sugar moiety or alpha-phosphate.

10 Thus, in another aspect the present invention provides a method of sequencing a target polynucleotide comprising the steps of:

- (a) extending a primer annealed to said target polynucleotide utilising a labelled nucleotide wherein the label is directly attached to the nucleotide;
  - (b) detecting the presence or absence of said labelled nucleotide within said  
15 extended primer;
  - (c) removal of said labelled nucleotide, and replacement of said labelled nucleotide with a degradation resistant nucleotide; and
  - (d) repeating steps a-c;
- wherein the 3' end of said primer comprises at least one degradation resistant  
20 nucleotide.

When the nucleotide is attached through a direct covalent bond to the label, then the replacement of the labelled nucleotide with an unlabelled nucleotide comprises removing the entire labelled nucleotide, and replacing it with a degradation resistant  
25 unlabelled nucleotide. "Degradation resistant" nucleotides are nucleotides which are not removed from the synthesised nucleotide sequence by degradation agents.

In one embodiment this is preferably done by removing nucleotides from the 3' end. Such agents include exonucleases, such as exonuclease III and includes DNA polymerases which possess 3' – 5' exonuclease activity. These enzymes include T4 polymerase, and *E.coli* DNA polymerase I (DNAPI). Preferably the degradation resistant nucleotides are alpha - s nucleotides. They may also be methylphosphonate linkages. Also for example boranophosphate modification at the alpha-phosphate group in 2'-deoxycytidine 5'-triphosphate (dCTP) (He *et al* 1999). Degradation reagents may also be chemical reagents such as mild reducing agent or mild acid. Physical degradation reagents include ultra-violet light for cleaving photocleavable bonds.

This embodiment is similar to that disclosed in US2003013101 which uses exonuclease activity to remove a nucleotide in its entirety, but in the present invention by using degradation resistant nucleotides the removal of more nucleotides than the last one added in a specific synthesis cycle is prevented. The incorporation of the degradation resistant nucleotide shifts the sequence register to the next position for the next cycle. The scheme also begins with primers that have degradation resistant nucleotides at their 3' end or when synthesis is initiated from a nick in double stranded DNA the first addition is of a degradation resistant nucleotide rather than a labelled nucleotide.

In another embodiment, one which can be carried out during either RNA (Gueroui 2002) or DNA synthesis, the degradation reagent cleaves an internal internucleoside bond. In one such preferred embodiment, following detection, a labelled degradation labile nucleotide is replaced by a degradation resistant nucleotide in order to shift the register to the next position in the sequence. In this case the degradation labile reagent may be a nucleotide modified at the 5' position with NH<sub>2</sub> (Wolfe 2003; Shchepinov 2001 and the degradation resistant nucleotide can be a normal nucleotide. NH<sub>2</sub> modified nucleotide analogs can be efficiently incorporated into DNA by the Klenow fragment of *Escherichia coli* DNA polymerase. Mild acid treatment of the resulting DNA specifically cleaves the modified internucleoside bond. This approach can be carried out by primer mediated DNA synthesis or promoter mediated RNA synthesis.

Details of synthesis of  $\text{NH}_2$  nucleotides is provided by Wolfe et al (2003). The nucleotides can be labelled by standard methods (e.g. see Hermanson, GT or Mitra 2003)

---

5 The labelled nucleotides that are incorporated may be dNTPs or ddNTPS. The disadvantage of adding dNTPS is there is no absolute certainty as to how many fluorescent bases become incorporated at each cycle. The disadvantage of labelled ddNTPS is that although only one labelled base will be added, after detection and removal of this base, the nucleotide which replaces it needs to allow incorporation of  
10 the next fluorescently labelled nucleotide and therefore cannot be blocked at the 3' end. However this may allow multiple bases to be incorporated. Therefore it is possible that more than one base may be added. This would shift the register beyond the last base that has been sequenced. This can be overcome by manipulating conditions, reaction components and reaction times so that, statistically, there is only  
15 the chance of incorporating one nucleotide complementary to each single molecule template, in one cycle. A pulse-chase reaction can be done in which the nucleotides are added for a short burst followed by a chase with apyrase enzyme which degrades free nucleotides. Addition of a labelled nucleotide may involve a different enzyme to addition of a degradation resistant nucleotide. Alternatively, removable blocking  
20 groups, as described in any of the prior art can be added to the 3' end of the unlabelled nucleotides. These can be removed before addition of the next labelled base.

25 Alternatively instead of using blocked fluorescent nucleotides or blocked degradation resistant nucleotides the incorporation of more than one nucleotide is allowed to occur. The number of fluorescent bases added can then be deduced from the

Also to prevent addition of more than one nucleotide at a time both the degradation resistant and degradation labile nucleotide may be modified or the reaction configured to prevent the addition of more than one nucleotide during one cycle.

5

In one preferred embodiment, a blocking group is added to the at the 3' position on the degradation labile nucleotide, restricting polymerization to a single base addition. The degradation reagent is able to remove the modified degradation labile nucleotide. After cleavage of this nucleotide by the degradation reagent, the degradation resistant nucleotide that replaces it also contains a blocking group at the 3' position, again restricting polymerization to a single base addition. This ensures that the sequence register is shifted by the required single position only. Although further degradation resistant nucleotides are not able to react at this 3' position, the degradation labile nucleotide is able to react at this position. Hence the process can continue. For example the degradation labile nucleotide may have an  $\text{NH}_2$  group at the 5' position which will have different reactivity than a 5' phosphate on the degradation resistant nucleotide.

10

15

Alternatively, although the degradation resistant nucleotide is not chemically blocked, the polymerizing reagent used to add the degradation resistant nucleotide may not be able to incorporate more than one degradation resistant nucleotides but the polymerizing reagent used for addition of the next labelled degradation labile nucleotide is able to add on to this base. Different enzymes have different processivities and different capacities to deal with natural and modified DNA nucleotides. The different steps may utilise different degradation reagents and different polymerization reagents. It should be noted that that although a frequent changing of polymerizing reagent is expensive, it is justifiable if sequencing is done on a large number of molecules in parallel. If the reactions are done in microfluidic channels the amount of reagents will be small and if a system of valves is incorporated onto a sequencing chip, the reagents which will usually be provided in excess amounts can be stored in designated chambers on the chip and re-used.

25

30

In one embodiment the label is attached to the nucleotide by means of a linker which comprises a cleavable bond, as described above. In the method of the invention, the replacement of said labelled nucleotide with an unlabelled nucleotide includes the removal of the label from said nucleotide.

5

The cleavable bond can be cleaved following the detection of the presence or absence of the labelled nucleotide. The label can be attached in such a way that it blocks the incorporation of further nucleotides. This ensures that the only one labelled nucleotide is incorporated.

10

The cleavable bond can be cleaved by means of light (if it is photocleavable) or by using a mild reducing agent if it comprises a disulphide bridge. Mild reducing agents include mercaptoethanol and DTT.

15

Thus in one aspect the present invention provides a method of sequencing a target polynucleotide comprising the steps:

- (a) Carrying out template derived nucleotide synthesis utilising a labelled nucleotide wherein the label is attached to the nucleotide via a cleavable linkage;
- (b) detecting the presence or absence of said labelled nucleotide within the synthesised polynucleotide;
- (c) cleaving said label from said nucleotide; and
- (d) repeating steps a-c.

20

25

In one preferred embodiment, the linkage attaching the label to the nucleotide comprises a binding pair. One member of the binding pair is linked to the nucleotide

therefore complementary to a particular spatial and polar organisation of the other member of the pair of molecules. Thus, the members of the pair have the property of binding specifically to each other. Examples of types of binding pairs are antigen-antibody, biotin-avidin, hormone-hormone receptor, receptor-ligand, enzyme-substrate. The use of a linkage comprising a binding pair means that the nucleotide added onto the primer may be labelled after it has been incorporated into the primer. The nucleotide is attached, preferably via a cleavable linker to one member of a binding pair. The label is attached to the other member of the binding pair. The label can then be attached indirectly to the nucleotide as the two members of the binding pair bind one another.

Each of the four types of nucleotides can be attached to a different binding pair member. The other members of the binding pair can be labelled differentially, i.e. each is attached to a different fluorophore or nanoparticle. This allows all of the nucleotides to be added at the same time. The nucleotide incorporated is then labelled with the respective fluorophore via the binding pair mechanism. For example adenine is attached to biotin, and cytosine is attached to digoxigenin. The fluorophore indicating the presence of adenine is attached to avidin, and that for cytosine is attached to anti-digoxigenin.

Thus in one aspect the present invention provides a method of sequencing a target polynucleotide comprising the steps of;

- (a) carrying out template derived nucleotide synthesis using a new nucleotide attached by a cleavable linkage to one member of a binding pair;
- (b) contacting said nucleotide with a label attached to the other member of the binding pair under conditions such that the two members of the binding pair bind to one another;
- (c) detecting the presence or absence of said label;
- (d) removal of said label and said binding pair by cleaving said cleavable linkage between the first and second members of the binding pair; and
- (e) repeating steps a-d.

Alternatively the fluorescent tag may comprise a nanoparticle.

The term "nanoparticle" as used herein means an individual particle which has a maximum dimension in any one direction of less than a micron. The nanoparticles of use in the present invention are preferably spherical, and/or preferably have a diameter of 20nm or less.

The fluorescent nanoparticle can be a particle which has a large number of fluorophores embedded within or on its' surface (eg latex particles). Alternatively, fluorescence emission or modulation may be an inherent property of the particle as is the case for semiconductor nanocrystals (Quantum Dot Corp., USA; Evident Technologies, USA), gold nanoparticles (Nanoprobes Inc., USA), plasmon resonant particles (PRPs) (Seashell Technologies, USA), Resonance light-scattering particles (RLP) or  $\text{TiO}_2$  nanoparticles (Paunesku 2003). Depending upon their size and material, semiconductor nanocrystals emit in different regions of the electromagnetic spectrum, even when excited with the same wavelength. Special coating procedures are applied to stabilize them in solution and make possible their conjugation with different objects. The advantage of nanocrystals is their high brightness of emission, high stability against photobleaching and that they have narrow emission spectrums, which facilitates multi-plexing. Semiconductor nanocrystals, of various emission wavelengths, with surfaces coated with streptavidin or biotin are available from Quantum Dot Corp. The streptavidin-biotin interaction can be mediated in the vendor supplied incubation buffer.

Nanoparticles can also be attached to a modified nucleotide via a thiol (sulfhydryl/mercaptan) group. Thiol groups can be attached to metals, in particular, gold. Alternatively, a linker can be used to attach the thiol to the nucleotide. The

is then deprotonated with 4-(dimethylamino)pyridine, rendering the nanocrystal water soluble, and facilitating reaction with thiolated nucleotides. After DNA modification, the particles are separated from unreacted DNA by dialysis or ultracentrifugation. Solubilised nanocrystals, emitting at various wavelengths are commercially available e.g. from Evident Technologies. Under the relatively mild reaction conditions used for enzymatic extension there is no appreciable thiol exchange.

As an alternative to linking the nucleotide to a nanoparticle through a thiol group, the binding pair linkage system described above can be used. The nucleotides can be attached to one member of a binding pair (e.g. biotin) through a cleavable linker and the nanoparticle may be coated with the other half of the binding pair, e.g. streptavidin. A photocleavable-Biotin-NHS reagent is commercially available from AmberGen which can react with amine groups in the nucleotides. A SNHS-SS-BIOTIN is available from Uptima and Pierce Biotechnology (EZ-Link Sulfo-NHS-SS-Biotin) which can be attached to amines on DNA bases and is cleavable by a mild reducing agent. In addition Photoprobe biotin reagent is available from Vector laboratories which allows biotin to be linked to DNA by heat or UV exposure, with the option of a cleavable disulphide bridge within the linkage. Using such a procedure a fraction of the nucleotides may be inappropriately labelled, in a way that perturbs the correct function e.g. hydrogen bonding of the base whereas another fraction may be correctly labelled.

Nanoparticles used in this invention are seen as a diffraction limited point source of fluorescence just like a single fluorophore. The advantage over other single molecule sequencing by synthesis approaches is that the nanoparticle is much easier to detect than a single dye molecule. Therefore a low grade CCD camera can be used for their detection and illumination may be from a mercury arc lamp. Lower grade objective lenses may be used and oil immersion lenses are not necessary. Sophisticated set-ups for background elimination, such as evanescent wave illumination may not be needed. Hence, the detection device is less sophisticated and cheaper than the instrument required for single dye molecule detection. Because a nanoparticle is easily

distinguished from artefacts and over background, sample preparation is easier and less stringent.

5 Nanoparticles can be easily removed after detection. A gold or CdSe/ZnS Quantum dot can be removed by treatment with a mild reducing agent such as DTT or mercaptoethanol. The Au-S bond, although thermodynamically stable, is kinetically labile, leading to thiol exchange in the presence of appropriate thiol-containing molecules in solution, particularly at elevated temperatures. It is also possible to attach DNA to nanoparticles via a binding pair as described above, which would avoid the use of thiols altogether. The linker connecting the base to one of the binding pairs or the thiol may contain a cleavable bond such as a disulphide bond which can be removed using a mild reducing reagent.

15 The nanoparticle may bear a positive charge which can interact with the nucleotide (Nakao 2003). The charged nanoparticle can be displaced by another charged species, following detection.

20 There are two specific ways that nanoparticles can be used in the invention to achieve sequencing by synthesis. The first involves addition of nanoparticle in order to label a base after it has been incorporated.

Thus in one embodiment the synthesis involves incorporation of a labelled nucleotide comprising:-

25 a) Incorporation of an unlabelled nucleotide adapted for the attachment of a nanoparticle; and

The second method by which nanoparticles can be used involves incorporation of nucleotides to which nanoparticles are already tethered.

Thus in another embodiment the replacement of the labelled nucleotide, which is labelled by means of a nanoparticle, with an unlabelled nucleotide comprises removing the nanoparticle from said labelled nucleotide. The nanoparticle can be removed by cleaving the cleavable bond in the linker attaching the nanoparticle to the nucleotide. The cleavable bond may be cleaved by light if it is photocleavable, or by means of a mild reducing agent such as DTT or mercaptoethanol if it is a disulphide bond..

The use of nanoparticles means that additions of multiple nucleotides can be detected more easily because the increase in signal is not expected to be quenched and so a digital increase in signal intensity can be expected with increasing number of nucleotides.

WO96/27025 discloses labelling nucleotides with microscopic beads in the context of sequencing by synthesis. The reagents used in the present invention are specifically nanoparticles of 20nm diameter and less as significantly larger beads would be too bulky to efficiently carry out the required molecular processes.

The nanoparticle strategy of the present invention differs from WO96/27025 in that the nanoparticles are not only used for labelling but may also be used to prevent the incorporation of more than one nucleotide at a time.

In addition to detection of nanoparticles due to fluorescence, they can also be detected efficiently by electron microscopies or scanning probe microscopies (e.g. see Csaki . 2001)

Nanoparticles can also be used to prevent a second base from being added by steric hindrance or repulsion. The nanoparticles may have a polarity, which repels another.

For example they may be positively or negatively charged (Nakao 2003) or they may have a magnetic polarity or spin (Lee 2003).

5 The methods of the present invention can be carried out on an array. The term "array" as used herein relates to a spatially defined arrangement of one or more nucleotides in a pattern on a surface. The array can consist of individual nucleotides present at at least 96, 384, 536, 10,000, 48,000 or 192,000 discrete locations on a surface. The array is preferably formed on a chip.

10 The array can be a random array wherein the nucleotides are attached to the surface randomly. Alternatively the arrays can be spatially addressed. The nucleotides can be arranged in a grid pattern, with regular spacing between each nucleotide. The nucleotides can be located in a "spot" along with a plurality of other nucleotides of the same sequence. Alternatively the arrays can comprise DNA colonies.

15 The polynucleotides can be attached either directly or indirectly to the surface. For example an enzyme, such as a ligase or polymerase, utilised in the process can be attached to a solid surface. The enzyme then binds the target polynucleotide, thus anchoring it to the solid surface.

20 Alternatively the polynucleotides can be captured by oligonucleotides which are attached to the surface. The capture can be by hybridisation of a single stranded oligonucleotide to a single stranded target or a single stranded region of a double stranded target. Alternatively, the polynucleotide or the surface immobilised capture probe, may comprise a sticky end or both may have a sticky end. The template and  
25 synthesized strand can be permanently linked to the surface by a ligation

nucleotides to form an array, and attaching nucleotides to an array are described in WO02/061126 and WO01/57248.

5 The surface is preferably glass, silica or a polymer such as PMDS. The substrate is preferably a glass slide, coverslip, silicon wafer, microfabricated chip or multi-well plate, such as a flat bottomed optical grade 96 well plate. The polynucleotides may be attached to material that coats the surface. For example aminosilane coated surfaces supplied by Corning Inc (USA) or Asper Biotech (Estonia) can be used. The surface may be coated with a gel material including a sol-gel. The polynucleotides may be  
10 attached to beads, particles, or structures such as nanobars or nanorods which may contribute to the generation or modulation of a FRET signal. The surface may be metallized with for example silver or gold particles to enhance a fluorescent or a raman signal (Malicka 2003; Kneipp 1999).

15 In addition, the surface or particles thereon may carry charge or be electrically biased or may be heated to control the sequencing process (Hamad-Schifferli, 2002). A charged surface is particularly useful to prevent non-specific interactions of nucleotides on the surface. Appropriate surface coatings include Polyethylamine as described by Braslavsky, 2003 and the DNA-bind slide available from VBC  
20 Genomics (Austria). An electric field generated at the surface is a useful way for controlling the attraction and repulsion of nucleotides at the surface (Asanov 1998; Sosnowski 1997)

25 Compared to the degree of parallelism currently available (96 Sanger sequencing reactions within individual capillaries on a state-of-the-art DNA sequencer), a whole wafer high-density oligonucleotide array has the capacity to analyse 60 million reactions (e.g. see [www.perlegen.com](http://www.perlegen.com) website). Until recently, the high cost of making individual photolithography masks meant that methods for making high-  
30 density oligonucleotide arrays were only available for mass production of arrays and were not accessible for the individual design of single arrays. However, the application of digital micromirror technology to array synthesis has made it much

more straightforward and cheaper to specify individual arrays. A fully integrated benchtop device for making, hybridising to and analysing high-density arrays can streamline the entire array experiment to within one day, (e.g. Geniom one; [www.febit.de](http://www.febit.de)).

5

The current technology (e.g. Geniom One) uses digital micro-mechanical mirrors to create an array by the spatially-selective deprotection of photolabile protecting groups on DNA chains growing on a surface. Each new array design can simply and rapidly be specified by software and there is no need to make photolithography masks. Arrays can be made such that the sequencing can be initiated either with an array of

— oligonucleotides directed to specific regions in the genome or with an array of n-mers (Gunderson et al, 1998) which will initiate the process at any position which seeds hybridisation. Presently, the standard software is configured to synthesize 48,000 oligonucleotides. However, by bypassing the software interface, it is possible to

15 synthesize at least 192,000 sequences on one chip in one synthesis run. All the synthesis, hybridisation and washing steps can be undertaken within the microfluidic channels of the chip provided by the manufacturer.

10

15

20

25

In one embodiment the method is carried out using an array wherein multiple copies of one primer are located within a localised area. The combined signal from all the nucleotides incorporated is detected i.e. the "bulk" signal is detected. The signal detected will be that relating to the nucleotide that is incorporated the most. In the embodiment wherein the nucleotides are all labelled identically and added individually, the strongest signal will be obtained when the nucleotide which is the correct base pair corresponding to the next base in the template is used. In the embodiment wherein the nucleotides are differentially labelled the fluorescent label

sequence of the template with the nucleotides having the same fluorescent label and the nucleotides having the same fluorescent label are added to the template and the nucleotides having the same fluorescent label are added to the template and the nucleotides having the same fluorescent label are added to the template.

Alternatively Single DNA molecule imaging can be used to detect the template and/or incorporations as a point-source of fluorescence, for each individual molecule.

5 Molecules within arrays are distributed at a density at which substantially all molecules are separated by a distance greater than that required for resolving them as separate entities (defined by the diffraction limit of light). Then, instead of analysing a single intensity value due to the combined signal from thousands of molecules, a "digital" signal from each molecule can be individually assessed. This enables  
10 heterogeneous reactions within a microarray spot, which would ordinarily be masked by the signal averaging of ensemble methods, to be detected.

Determining which nucleotide has been incorporated in a single molecule allows multiple copies of a polynucleotide to be sequenced individually, in parallel. The  
15 present invention enables, for example around  $10^3$  sequence passes within a microarray spot. If a thousand copies of a polynucleotide can be sequenced at the same time, it is effectively equivalent to repeating the sequencing a thousand times. This considerably reduces the amount of time required to carry out this work, as compared to the traditional Sanger dideoxy techniques. It also provides increased confidence levels.  
20 This method eliminates the need for costly amplification steps, and can be used to provide haplotype information.

The single molecule sequencing approach developed previously involves a "random" display of the diversity of molecules to be sequenced without any deliberate  
25 organization of the molecules by spatially addressable arrays. The methods of the invention can be applied to such types of random arrays of single molecules. In such a set-up although there may be several other copies of the same sequence present elsewhere, at undefined positions on the surface it will be difficult to extract statistical confidence in a sequence due to heterozygosity and the presence of other closely  
30 related sequences. Therefore, each molecule is essentially sequenced with one pass only. By contrast single molecule sequencing within spatially addressable microarrays enables for example, around  $10^3$  molecules of the same species to be sequenced

within each array spot. Hence, the confidence levels with which the sequence will be obtained will be unprecedented. If related but different sequences are captured within a spot, their identity will become apparent after several cycles of base addition. If the sample is heterozygous, then the presence of two species within the spot will be seen.

5

Although the methods of the present invention are preferably carried out on a solid surface, they can also be conducted on molecules which are free in solution for example in the wells of a microtitre plate or within micro or nano-scale vials, wells or structures (Levene et al 2003).

10

The methods of this invention are preferably undertaken on surfaces because it is easier to organise and/or monitor reactions on a surface than reactions freely diffusing in solution. However, when the molecules are immobilised on beads which are able to diffuse, one can take advantage of the improved reaction kinetics of solution phase reactions. A sequence strategy has been described for molecules immobilised on beads (Brenner et al 1999) and the methods of the present invention could be applied on this platform.

15

As the invention can be applied in a single molecule detection mode it is very sensitive and can be performed on small amounts of sample material. Hence the invention can also be applied in a context where one or very few molecules are available, such as from ancient DNA or a forensic sample.

20

Although the invention can be carried out on a purified fragment of a polynucleotide, it offers particular advantages for sequencing polynucleotides directly from a complex mixture such as sheared/fragmented genomic DNA, a mRNA population or a

25

in different fractions. This is needed particularly when the invention is implemented in bulk mode. When the invention is implemented in single molecule mode the identity of each individual molecule is fully determined regardless and hence a contaminating sequence can be identified. However to avoid unnecessary sequencing repetitive DNA can be suppressed by for example subtraction with Cot-1 DNA. The invention can be implemented on spatially addressable arrays so that different regions of the genome or different species in a mRNA population are captured at specific known locations. One advantage of this is that capture probes provide a certain length of sequence information even before the sequencing by synthesis data is obtained.

Methods of this invention can be carried out in heterogeneous way, where reaction components for the different steps of the reaction are provided at separate stages. The methods can also be carried out in a "homogeneous" way, where all the components required for the reaction are provided in the reaction vessel from the start. Then cyclical electromagnetic modulation, for example for cleaving a linkage provides a clocking mechanism for shifting the sequence register. Furthermore, some of the methods of this invention can be carried out in real-time, by providing reaction components and then continuously monitoring the reaction. Preferably for this embodiment the signal is detected by a FRET mechanism as described below.

### Sequencing by FRET

In fluorescence resonance energy transfer (FRET), a donor fluorophore molecule absorbs excitation energy and delivers this via dipole-dipole interaction to a nearby acceptor fluorophore molecule (Stryer, L. and Haugland, R.P. 1967.). Fluorescence resonance energy transfer can be used to cut out background fluorescence in single molecule experiments (Braslavsky *et al* 2003). Recently, a new way of using FRET in a DNA assay, termed iFRET has been introduced in which the donor dye is an DNA intercalating dye that is used to stain DNA (Howell WM *et al.* 2002). iFRET is reported to give fluorescence values that are 2.5 times greater than those obtained from the intercalating dye alone, and more than 40 times greater than those from conventional FRET. It is suggested that the reason for the difference may be that the iFRET system involves the channelling of an accumulation of energy from a chain of

donor dye molecules (in contrast to a single donor in the FRET system) into the acceptor moiety, which is then able to re-emit energy unhindered. Double-strand, DNA-specific intercalating dye (e.g., SYBR Green I) has been used as a FRET donor, with a conventional FRET acceptor.

5

A FRET mechanism can be implemented with the sequencing by synthesis methods described in this invention. One embodiment of the present invention involves the detection at the single molecule level, using FRET between two or more FRET partners. The FRET partnership system comprises two or more partners each attached to a reaction component selected from the group comprising nucleotide, the template, the polymerasing agent or any other reagent involved in the polymerization reaction. Donor-acceptor fluorophore pairs are chosen so that the emission spectrum of the donor overlaps with the excitation spectrum of the acceptor; many different combinations of available fluorescent labels can be used.

10

15

In one preferred embodiment the FRET means of detection is utilised in a method wherein the labelled nucleotide is detected as it approaches the target DNA molecule. As the labelled nucleotide is brought into the proximity of the target polynucleotide during polymerisation, the FRET reaction occurs between the label on the nucleotide and a FRET partner. This reaction can be detected. The FRET label is attached to the nucleotide through the beta or gamma phosphate groups. These phosphate groups are removed as the nucleotide is added during extension, so effectively the detection of the label, the extension, and the replacement of the labelled nucleotide with an unlabelled nucleotide occur almost simultaneously. When the nucleotide has been incorporated it is no longer labelled. The released pyrophosphate is free to diffuse out of FRET range.

20

25

In another preferred embodiment FRET occurs between a DNA stain (e.g. an intercalating dye) bound to DNA and one or more FRET partners attached to another polymerisation reaction component such as the nucleotide or polymerising reagent.

5 The bound DNA stain may act as FRET donor or acceptor. It is simple to add a DNA stain that incorporates at multiple positions along a template molecule so that it can contribute to a FRET reaction anywhere along the extending chain. One of the FRET partners may be the fluorescently labelled nucleotide, which is utilised to extend the polynucleotide being synthesised. The fluorescent label may be directly or indirectly attached to nucleotide, and it may be a nanoparticle. Preferably the DNA stain is not  
10 the first FRET Donor as this could lead to it's wholesale photobleaching; although this can be minimised with judicious choice of antifade composition. Several DNA stains are available for staining double-stranded DNA and a few of these are also able to stain single-stranded DNAs relatively efficiently, e.g. SYBR Gold.

15 Alternatively, or additionally the FRET partner can be attached to the polymerase, for example the DNA polymerase. The FRET label may be in the form of a semiconductor nanocrystal/Quantum Dot, as these do not photobleach which is important as it is desirable to retain the same polymerase throughout synthesis.

20 Multiple FRET interactions can take place when the excitation and emission spectrum for FRET partners overlap. The first FRET partner is excited at one wavelength, and it's emission wavelength overlaps with the excitation wavelength for the second a FRET partner. The second FRET partner has an emission wavelength which overlaps with the excitation wavelength for a third FRET partner. In this way a  
25 chain of energy transfers can take place, when the FRET partners are within FRET range and the first donor has been excited. This can result in a large stokes shift i.e. large separation of excitation from emission. This allows the signal to be read at a wavelength far removed from to the original excitation wavelength, which is advantageous for eliminating bleed-through from the excitation source into the  
30 detection channel. Importantly, this method also ensures that all of the components (the target polynucleotide, the labelled nucleotide and the polymerase) are all in close proximity. In some instances an anti-stokes shift may be utilised.

Howell *et al* describe a system in which the intercalating dye acts as donor. A Single molecule-system may involve syber-green-1 as the donor and a rox-labelled nucleotide as the acceptor. As Quantum Dots can be excited at various wavelengths, when they are used as the acceptor, donors emitting at various wavelengths can be used, e.g. DAPI or SYBR gold. Alternatively, the incorporated fluorescent nucleotide or a fluorescent nanoparticle can act as the donor and an intercalating dye such as PO-PO3 can be used as the acceptor (Nakayama et al 2003). The Quantum Dot can be excited at a wavelength far removed from the acceptor dye. The signal produced would be due to the localised excitation by the Quantum Dot of a few fluorescent dyes in its locality. Following detection of the FRET signals, an image of the polynucleotide polymer can be taken by exciting the DNA stain directly. The FRET signal can then be superimposed on the polynucleotide polymer image, to determine where incorporation has occurred.

Because energy transfer to the acceptor is from a highly localised source, background fluorescence from anything beyond the FRET range, which is about 10nm, does not contribute to background fluorescence. Hence FRET would enable reactions to be monitored continuously without the need for washing away of unbound fluorescent dyes or nanoparticles. This would enable addition of more than one nucleotide to be detected in real-time. The system could be homogeneous in that all that was needed for the reaction can be placed in the reaction vessel at the start of synthesis. It would be desirable to retain some form of agitation or mixing of the reaction solution to enable pyrophosphate to diffuse out of FRET range after it has been released.

In accordance with the above in one aspect the invention provides a method of

Preferably this method is used to carry out real time monitoring of the sequence.

The polymerisation reaction components include the polymerizing reagent and the template polynucleotide. Preferably a DNA stain is used to label the template and a  
5 Quantum Dot is used to label the polymerizing agent.

A nucleotide that may be temporarily resident within the FRET range of a polymerizing agent or a template molecule, may or may not get incorporated depending on whether it is the correctly matched nucleotide for the position in  
10 question. This temporary resident of the FRET locality must be distinguished from a nucleotide that is actually incorporated. This can be done by utilising information gathered prior to the reaction about, for example, the longevity or strength of the FRET signal depending on whether it originates from a nucleotide temporarily  
15 resident within the FRET locality or a properly incorporated nucleotide. WO00/36151 describes a mechanism in which the dye attached to the nucleotide remains quenched by a quencher until incorporation of the nucleotide occurs at which point the quencher becomes detached and allows the dye to fluoresce freely. The drawback of this approach is that there is likely to be loss of quenching which is not due to loss of  
20 quencher but is due to thermal or structural fluctuations or photobleaching. An alternative way of measuring incorporation in the context of the present invention is by detecting quenching/de-quenching or preferably a wavelength shift with a FRET partner which occupies a different reaction component than the nucleotide itself. For example, the emission due to the FRET partner on the template may be modified by a FRET Partner on the beta or gamma phosphate of the nucleotide. When the nucleotide  
25 is incorporated and pyrophosphate is released, the FRET interaction is abolished and hence a fluorescence property of the FRET partner on the template is modified, eg, it emits fluorescence at a shifted wavelength. The first donor in this scheme may be a Quantum Dot attached to the polymerizing agent and the whole process may be designed to have multiple FRET interactions which are able to be monitored in real  
30 time, by using for example an image splitter such as the Quad-view from Optical Insights Inc. (USA).

Appropriate anti-fades can be used to attenuate photobleaching. This can include the provision of Oxygen scavengers and reducing agents such as DTT and Mercaptoethanol.

## 5 Linear Polymer Display

Genomic sequence would have much greater utility if haplotype information (the association of alleles along a single DNA molecule derived from a single parental chromosome) could be obtained over a long range. This is possible by combining the SbS process of the present invention with the single molecule display of linearised genomic DNA described in WO02/074988. Here each template molecule is sorted on the array, and combed out to provide a linear display of sequence along its length.

Polymerisation can then be seeded at multiple positions on each linear molecule, e.g. optimised to be every 10kb apart. The incorporations are monitored as slowly migrating point sources of fluorescence along the linearised DNA polymer. The introduction of nicks in the double stranded DNA is sufficient to prime synthesis. DNA stains such as SYBR Gold can be used to visualize the DNA polymer. This method can be used to confirm that it is the template polynucleotide that is being sequenced, as opposed to some other contaminant.

## 20 Resolving ambiguities in the sequence

If the sequencing approaches are carried out in a microarray format and the array making and sequencing is iterated, then ambiguities in the number of bases at any particular position can be resolved by making probes that would address each of the suspected sequence possibilities in the next array synthesis.

The results can be displayed with confidence levels for each base and where bases

### Sequencing to obtain gene expression information

The method of the present invention can be adapted to obtain gene expression data, particularly from a single cell. Once a certain length of sequence information has been obtained, it can be used to identify the mRNA species. The method can be modified for sequencing mRNA. Thus in one embodiment the target polynucleotide comprises mRNA. The mRNA can be hybridised to primers which are designed to hybridise to any mRNA molecule. For example, primers can be designed to hybridise to all sample mRNA species at a specific point in the mRNA primary structure. This point could be the polyadenylation signal, AAUAAA, the Poly A tail at the 3' end or the cap structure at the 5' end or a specific sequence clamped onto the 5' or 3' end. Preferably the primers are attached to a solid surface, and more preferably form an array.

- Thus in one aspect the present invention provides a method of sequencing mRNA comprising:
- a) contacting an array of probes designed to hybridise to mRNA molecules with a sample of mRNA under conditions whereby the mRNA will hybridise to said probes; and
  - b) sequencing said mRNA utilising a method as described herein

### Co-sequencing two samples to find the differences between them

The DNA or mRNA from two or more individuals or populations can be compared by differentially labelling each template (i.e. labelling the template with a different label for each population or individual), immobilising them on a surface and then sequencing them simultaneously. The templates can be labelled, for example by attaching an oligonucleotide containing different fluorescent dyes by using RNA ligase.

- The templates can be immobilised by attaching the labelled nucleotides to a surface as described above. The templates can be used to form an array. Alternatively the templates can be captured on to an array. This can be done for example, if the

templates nucleotides are allowed to hybridise to primers which themselves form an array.

5 After, the templates are immobilised on the surface, the foci for each sample will be detected and recorded. Following this, the label can be photobleached and the sequencing can commence.

#### Detection schemes and instrumentation

10 The images of the polynucleotides are projected onto the array of a Charge-couple device (CCD) camera, from which they are digitized and stored in memory. The images stored in memory are then subjected to image analysis algorithms. These algorithms can distinguish signal from background, monitor changes in signal characteristics, and perform other signal processing functions. The memory and signal processing may be performed off-line on a computer, or in specialized digital signal  
15 processing (DSP) circuits controlled by a microprocessor.

When the base-by-base incorporation of labelled nucleotides is monitored on molecules in bulk (i.e. the combined signal from a population of polynucleotides within one spot on an array is measured), the established methods for scanning  
20 microarrays can be used to monitor incorporation at microarray spots after each base addition. For example a Genepix scanner (Axon instruments) or a Scanarray (Packard) which can be linked to four different laser lines can be used.

25 When individual molecules within the microarray spot are analysed directly, then wide field CCD imaging is used. CCD imaging enables a population of single molecules distributed 2-dimensionally on a surface to be viewed simultaneously.

magnification of the objective, any magnification due to the C-mount and, the size and number of pixels of the CCD chip. Typically, a microarray spot can be viewed by either a 40X or 60X objective depending on CCD camera and C-mount. Therefore to view large regions of a slide (several cm<sup>2</sup>) multiple images must be taken. A low noise high sensitivity camera is used to capture images. There are several camera models that can be used; Cooled Micromax camera (Roper scientific) controlled by MetaMorph (also MetaView software; both from Universal Imaging). MetaMorph can be run on a Dell OptiPlex GX260 personal Computer.

**10     Microarray spot-finding and single molecule imaging within microarray spots**  
MetaMorph's optional microarray module and a low magnification objective can be used to locate spots before taking a CCD image of each of the spots using higher magnification.

15     As the signal from the spots containing singly resolvable molecules is very low under low magnification, a marker dye, which emits at a different wavelength to the sample emission can be included in the spots to help locate them. The objectives need to be of high numerical aperture (NA) in order to obtain good resolution and contrast. The integration of an autofocus capability within the procedure to maintain focus as the slide is scanned, is useful especially when Total Internal Reflection Fluorescence microscopy (TIRF) is employed. Software can be used to control Z movement (integral to motorized microscopes) for the purpose of autofocus. Images of microarray spots can be obtained by x-y movements of the sample stage (e.g. using Prior Scientific's Proscan stage under MetaMorph control). To avoid photobleaching it is advisable to use a shutter (e.g. from Prior Scientific) to block off illumination while moving from one spot to another. A controller can be used to control X-Y stage, the filter wheels and shutter, (eg Prior Scientific ProScan).

30     Once the spots are found, their coordinates are recorded by the software controlling the instrument and then after each base addition, a CCD image is taken of each spot of the microarray.

In addition to the instrument being used for looking at a microarray where template molecules have been captured by probes, a large number of samples can be gridded (as a microarray) and then the instrument can be used to analyse each spot. The samples may be individual nucleotide populations or a set of differentially labeled nucleotide populations.

Two imaging set ups, Total Internal Reflection Fluorescence microscopy (TIRF) and epi-fluorescence microscopy have been used.

#### 10 **Epifluorescence Microscopy**

Images of single molecules labeled with a single dye molecule can be obtained using a standard epi-fluorescence microscopy set up, using high numerical aperture (NA) objectives and a high grade CCD camera. However, the image can be hazy. In order to obtain a clearer image it is preferable to use deconvolution software to remove the haze. Deconvolution modules are available as drop-ins for MetaMorph software. When the single molecules are labeled with nanoparticles the camera and objectives may be of a lower grade and oil-immersion objectives may not be required.

#### **Total Internal Reflection Fluorescence Microscopy (TIRF)**

20 TIRF enables very clean images to be obtained, for example using off the shelf system for Objective style TIRF (such as those produced by Olympus or Nikon). A full description can be found in the brochure at the following website: [www.nikon-instruments.com/uk/pdf/brochure-tirf.pdf](http://www.nikon-instruments.com/uk/pdf/brochure-tirf.pdf)

25 Objective style TIRF can be used when the sample is on a coverslip. However, it is not compatible when the sample is on a microarray slide. For this Prism type TIRF

**Multi-colour single molecule imaging**

When the sequencing strategy involves the sequential addition of each of the four nucleotides all labeled with a single fluorophore such as Cy3, then a single CCD image is taken after each base addition. However, if each nucleotide is differentially labeled (i.e. each nucleotide type is labelled with a different fluorophore) and added simultaneously, then the signal from each of the different fluorophores needs to be acquired distinguishably. This can be done by taking four separate images by switching excitation/emission filters. Alternatively, an image (Wavelength) splitter such as the Dual View (Optical Insights, Santa Fe, NM) or W View (Hamamatsu, Japan) which direct the light through two separate bandpass filters with little loss of light between them, can be used for imaging two different wavelengths onto different portions of a CCD chip. Alternatively the light can be split into four wavelengths and sent to the four quadrants of a CCD chip (e.g Quad view from Optical Insights). This obviates the need to switch filters using a filter wheel. A MetaMorph drop-in for single image dual or multi-emission optical splitters can also be employed. Image splitting can be used to monitor FRET.

**Monitoring sequencing on single molecules randomly distributed on a surface**

As an alternative to microarray spot finding prior to single molecule imaging and for implementations where the single molecules to be analysed are not organised within the spatially addressable microarray spots, a series of images of the surface can be taken by x-y translation of the slide. A super-wide field image is then composed by stitching each of the images together. This process can be automated to form a high throughput system, utilising computer software to control the process. Such a process could include robotic sample preparation. A small desktop apparatus with an integrated biosensor could be developed.

Where real-time sequencing is carried out, translation of sample with respect to CCD camera may be too slow to detect each molecular event. Therefore a method for collecting single molecule data on a surface by taking images simultaneously with an array of CCD chips can be applied.

This methods of this invention may aid functional genomics by enabling phenotype-  
~~genotype correlations to be made at a high resolution. Massively parallel sequencing~~  
would enable fine-scale genotype-phenotype correlations to be made for biological  
5 studies, improve methods for determining the functional contribution of genes to  
diseases and be the basis for developing personalized medicines which are tailored to  
a persons functional genetic make-up.

## REFERENCES

- 5 Asanov AN, Wilson WW, Oldham PB. Regenerable biosensor platform: a total internal reflection fluorescence cell with electrochemical control. *Anal Chem.* 1998 Mar 15;70(6):1156-63
- 10 Braslavsky I, Hebert B, Kartalov E, Quake SR. *Proc Natl Acad Sci U S A.* 100:3960-4. Sequence information can be obtained from single DNA molecules. (2003)
- 15 Brenner S, Johnson M, Bridgham J, Golda G, Lloyd DH, Johnson D, Luo S, McCurdy S, Foy M, Ewan M, Roth R, George D, Eletr S, Albrecht G, Vermaas E, Williams SR, Moon K, Burcham T, Pallas M, DuBridge RB, Kirchner J, Fearon K, Mao J, Corcoran K. Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. *Nat Biotechnol.* 2000 Jun;18(6):630-4
- 20 Csaki A, Moller R, Straube W, Kohler JM, Fritzsche W. DNA monolayer on gold substrates characterized by nanoparticle labeling and scanning force microscopy. *Nucleic Acids Res.* 2001 Aug 15;29(16):E81.
- 25 Czapinski, J. L.; Sheppard, T. L.; *J. Am. Chem. Soc.*; (Communication); 2001; 123(35); 8618-8619 Nucleic Acid Template-Directed Assembly of Metallosalen-DNA Conjugates
- 30 Delaney JC, Henderson PT, Helquist SA, Morales JC, Essigmann JM, Kool ET. High-fidelity in vivo replication of DNA base shape mimics without Watson-Crick hydrogen bonds. *Proc Natl Acad Sci U S A.* 2003 Apr 15;100(8):4469-73. Epub 2003 Apr 03.
- 30 Eckardt LH, Naumann K, Pankau WM, Rein M, Schweitzer M, Windhab N, von Kiedrowski G. DNA nanotechnology: Chemical copying of connectivity. *Nature.* 2002 Nov 21;420(6913):286

Ecker DJ, Vickers TA, Hanecak R, Driver V, Anderson K. Rational screening of

~~oligonucleotide combinatorial libraries for drug discovery. Nucleic Acids Res. 1993~~

Apr 25;21(8):1853-6). Finzi, L and Gelles J Science 267:378 (1995)

5

Hermanson GT, Bioconjugate Techniques, Academic Press 1996

Fodor S.P.A *et al* Science 251:767 (1991)

10

Gueroui Z, Place C, Freyssingeas E, Berge B. Observation by fluorescence microscopy of transcription on single combed DNA. Proc Natl Acad Sci U S A. 2002 Apr 30;99(9):6005-10.

15

Gunderson KL, Huang XC, Morris MS, Lipshutz RJ, Lockhart DJ, Chee MS. Mutation detection by ligation to complete n-mer DNA arrays. Genome Res. 1998 Nov;8(11):1142-53.

20

Hamad-Schifferli K, Schwartz JJ, Santos AT, Zhang S, Jacobson JM. Remote electronic control of DNA hybridization through inductive coupling to an attached metal nanocrystal antenna. Nature. 2002 Jan 10;415(6868):152-5

25

He K, Porter KW, Hasan A, Briley and JD, Shaw BR Synthesis of 5-substituted 2'-deoxycytidine 5'-(alpha-P-borano)triphosphates, their incorporation into DNA and effects on exonuclease.. Nucleic Acids Res. Apr 15;27(8):1788-94 (1999)

~~Hegner H *et al* FEBS Letters 336(2):452 (1993)~~

Kneipp K, Kneipp H, Itzkan I, Dasari RR, Feld MS.

Ultrasensitive chemical analysis by Raman spectroscopy.

Chem Rev. 1999 Oct 13;99(10):2957-76.

- 5 Kricka LJ. Stains, labels and detection strategies for nucleic acids assays. *Ann Clin Biochem.* 2002 Mar;39(Pt 2):114-29.

- 10 Lee HY, Sacho Y, Kanki T, Tanaka H, Shirakawa H, Cheon JW, Yoon JH, Kang NJ, Park JI, Kawai T. J DNA-directed magnetic network formations with ferromagnetic nanoparticles. *Nanosci Nanotechnol.* 2002 Dec;2(6):613-5.

AJ Lacey *Light Microscopy in Biology, A practical Approach.* OUP

- 15 Levene MJ, Korlach J, Turner SW, Foquet M, Craighead HG, Webb WW. Zero-mode waveguides for single-molecule analysis at high concentrations. *Science.* 2003 Jan 31;299(5607):682-6.

- 20 Malicka J, Gryczynski I, Fang J, Kusba J, Lakowicz JR. Increased resonance energy transfer between fluorophores bound to DNA in proximity to metallic silver particles. *Anal Biochem.* 2003 Apr 15;315(2):160-9.

- 25 Mitra RD, Shendure J, Olejnik J, Edyta-Krzyszanska-Olejnik, Church GM. Fluorescent in situ sequencing on polymerase colonies. *Anal Biochem.* 2003 Sep 1;320(1):55-65.

- 30 Nakayama H, Arakaki A, Maruyama K, Takeyama H, Matsunaga T. Single-nucleotide polymorphism analysis using fluorescence resonance energy transfer between DNA-labeling fluorophore, fluorescein isothiocyanate, and DNA intercalator, POPO-3, on bacterial magnetic particles. *Biotechnol Bioeng.* 2003 Oct 5;84(1):96-102.

Nakao Hidenobu, Hiroshi Shiigi, Yojiro Yamamoto, Shiho Tokonami, Tsutomu Nagaoka, Shigeru Sugiyama, and Toshio Ohtani. Highly Ordered Assemblies of Au Nanoparticles Organized on DNA. *Nano-letters* Vol 3 pp 1391–1394 2003

---

- 5 Paunesku T, Rajh T, Wiederrecht G, Maser J, Vogt S, Stojicevic N, Protic M, Lai B, Oryhon J, Thurnauer M, Woloschak G. Biology of TiO<sub>2</sub>-oligonucleotide nanocomposites. *Nat Mater.* 2003 May;2(5):343-6.
- 10 Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A.* 74:5463-7 (1977)
- 15 Shchepinov MS, Korshun VA. Design of multidye systems for FRET-based applications. *Nucleosides Nucleotides Nucleic Acids.* 2001 Apr-Jul;20(4-7):369-74.
- 20 Shchepinov, M.S., Denissenko, M.F., Smylie, K.J., Worl, R.J., Leppin, A.L., Cantor, C.R. and Rodi, C.P. (2001) Matrix-induced fragmentation of P3'-N5'-phosphoramidate-containing DNA: high-throughput MALDI-TOF analysis of genomic sequence polymorphisms. *Nucleic Acids Res.*, 29, 3864–3872
- 25 Sosnowski RG, Tu E, Butler WF, O'Connell JP, Heller MJ. Rapid determination of single base mismatch mutations in DNA hybrids by direct electric field control. *Proc Natl Acad Sci U S A.* 1997 Feb 18;94(4):1119-23
- 30 Stryer, L. and Haugland, R.P. 1967. Energy transfer: A spectroscopic ruler. *Proc. Natl. Acad. Sci.* 58: 719-726

Xu Y, Karalkar NB, Kool ET. Nonenzymatic autoligation in direct three-color detection of RNA and DNA point mutations. *Nat Biotechnol.* 2001 Feb;19(2):148-52.

Zander, C, Enderlain, J, Keller, RA, Singel Molecule Detection in Solution Wiley-VCH 2002.

## CLAIMS

1. A method of sequencing a target polynucleotide comprising the steps of:
- 5 (a) Carrying out template derived nucleotide synthesis utilising a labelled nucleotide;
- (b) detecting the presence or absence of said labelled nucleotide;
- (c) replacing said labelled nucleotide with an unlabelled nucleotide; and
- 10 (d) repeating steps a) to c)
- with the proviso that if said labelled nucleotide is labelled with a label directly attached to the nucleotide, then the replacement of said labelled nucleotide comprises removal of the whole of said labelled nucleotide and replacement with an unlabelled nucleotide, and only said labelled nucleotide can be removed.
- 15 2. A method as claimed in claim 1 wherein said target polynucleotide is attached to a solid surface.
3. A method as claimed in claim 1 or claim 2 wherein said labelled nucleotide is
- 20 labelled with a fluorescent tag.
4. A method as claimed in claim 3 wherein said fluorescent tag is attached directly to said nucleotide.
- 25 5. A method as claimed in claim 4 wherein step (c) comprises removal of said ~~labelled nucleotide and replacement with an unlabelled nucleotide wherein said~~

7. A method as claimed in claim 6 wherein said nanoparticle is a semiconductor nanocrystal.
8. A method as claimed in any one of claims 3, 6 or 7 wherein said fluorescent tag or said nanoparticle are attached to said labelled nucleotide by a linkage.
9. A method as claimed in claim 8 wherein said linkage comprises a cleavable bond.
10. A method as claimed in claim 8 or claim 9 wherein said linkage comprises a binding pair.
11. A method as claimed in any one of claims 7 to 10 wherein step (b) comprises incorporation of an unlabelled nucleotide adapted for the attachment of a nanoparticle; and attaching said nanoparticle to said unlabelled nucleotide.
12. A method as claimed in any one of claims 7 to 11 wherein step (c) comprises removing the nanoparticle from said labelled nucleotide.
13. A method as claimed in claim 12 wherein said nanoparticle is removed from said labelled nucleotide by cleaving the cleavable bond in the linkage attaching said nanoparticle to the nucleotide.
14. A method as claimed in claim 13 wherein said linkage attaches one member of a binding pair to the nucleotide, and the other member of the binding pair is attached to said nanoparticle.
15. A method as claimed in claim 6 wherein said linkage comprises a binding member attached by a cleavable bond to said nucleotide and the other binding member is attached to said fluorescent tag.
16. A method as claimed in claim 15 wherein step (d) comprises removal of said fluorescent tag by cleaving said cleavable bond.

17. A method as claimed in anyone of claims 1 to 3 wherein step (b) is carried out by means of an imaging technique utilising FRET (fluorescent resonance energy transfer).

18. A method as claimed in claim 17 wherein said target polynucleotide is treated with a DNA intercalating dye.

19. A method as claimed in claim 17 or claim 18 wherein said labelled nucleotide is labelled with a label that acts as a FRET partner.

20. A method as claimed in anyone of claims 17 to 19 wherein steps (a)-(c) occur simultaneously.

21. A method as claimed in any one of claims 1 to 20 wherein said target polynucleotide forms part of an array.

22. A method as claimed in claim 21 wherein step (c) comprises measuring the signal generated by a plurality of said labelled nucleotides.

23. A method as claimed in claim 21 wherein step (c) comprises detecting the presence or absence of said labelled nucleotide for each individual polynucleotide.

24. A method as claimed in claim 23 wherein said detection is carried out by means of single DNA molecule imaging.

27. A method as claimed in claim 26 wherein the label on said labelled nucleotide acts as a FRET partner to said DNA intercalating dye.

5 28. A method of comparing two or more polynucleotide sequences comprising:  
a) differentially labelling the nucleotide sequences being compared;  
b) immobilising said nucleotide sequences on a surface;  
c) detecting the locus of each nucleotide sequence; and  
d) sequencing said polynucleotide sequences using a method as claimed in any of  
10 claims 1 to 27.

29. A method as claimed in claim 28, further comprising photobleaching the label prior to the sequencing of said polynucleotide sequence.

15 30. A method of resolving ambiguities in a polynucleotide sequence comprising:  
a) identifying an area of ambiguity in a polynucleotide sequence;  
b) designing probes for each of the suspected sequence possibility; and  
c) utilising the primers formed to sequence said polynucleotide sequence  
utilising a method as claimed in any of claims 1 to 27.

20 31. A method of sequencing mRNA comprising:  
a) contacting an array of probes designed to hybridise to mRNA molecules with  
a sample of mRNA under conditions whereby the mRNA will hybridise to  
said probes; and  
25 b) sequencing said mRNA utilising a method as claimed in any one of claims 1 to  
27.

32. A method as claimed in 31 wherein said probe is designed to hybridise to the  
polyadenylation signal, 5' cap, 3' tail or the poly A tail.

30 33. A method of sequencing a target polynucleotide comprising the steps of:  
(a) treating said target polynucleotide with an intercalating dye;

(b) extending a primer annealed to said target polynucleotide utilising a nucleotide labelled with a label which acts as a FRET partner to said DNA intercalating dye;

(c) detecting the presence or absence of said nucleotide by means of an imaging technique that utilises FRET; and

(d) repeating steps a-c;

wherein steps (a) and (b) can occur in any order.

34. A method of sequencing a target polynucleotide comprising the steps of:

(a) extending a primer annealed to said target polynucleotide utilising a labelled nucleotide wherein the label is directly attached to the nucleotide;

(b) detecting the presence or absence of said labelled nucleotide within said extended primer;

(c) removal of said labelled nucleotide, and replacement of said labelled nucleotide with an unlabelled degradation resistant nucleotide; and

(d) repeating steps a-c;

wherein the 3' end of said primer comprises at least one degradation resistant nucleotide.

35. A method of sequencing a target polynucleotide comprising the steps:

(a) extending a primer annealed to said target polynucleotide utilising a labelled nucleotide wherein the label is attached to the nucleotide via a cleavable linkage;

(b) detecting the presence or absence of said labelled nucleotide within said extended primer;

(c) cleaving said label from said nucleotide; and

(b) contacting said nucleotide with a label attached to the other member of a binding pair under conditions such that the two members of the binding pair bind to one another;

(c) detecting the presence or absence of said label;

5 (d) removal of said label and said binding pair by cleaving said cleavable linkage;  
and

(e) repeating steps a-d.

37. A method of sequencing a target polynucleotide, comprising the steps of:

10 (a) carrying out template derived polynucleotide synthesis utilising a nucleotide  
labelled with a Fret partner and at least one other polymerisation reaction

... component labelled with a Fret partner;

(b) determining the nucleotide incorporated by detecting Fret interactions; and

(c) repeating steps (a) and (b).

15

20

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☒ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**